

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
31 May 2001 (31.05.2001)

PCT

(10) International Publication Number
WO 01/38546 A1

(51) International Patent Classification⁷: C12N 15/54, 9/12, 1/21, C12Q 1/68 // (C12N 1/21, C12R 1:19)

(21) International Application Number: PCT/US00/31830

(22) International Filing Date: 17 November 2000 (17.11.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data: 60/167,066 23 November 1999 (23.11.1999) US

(71) Applicant: AMERSHAM PHARMACIA BIOTECH, INC. [US/US]; 800 Centennial Avenue, Piscataway, NJ 08855-1327 (US).

(72) Inventors: EVANS, Steven, J.; 6 Layfield Road, Brunton Park, Gosforth, Newcastle-upon-Tyne NE3 5AA (GB). MAMONE, Anthony; 902 Holland Drive, Somerset, NJ 08873 (US). DAVIS, Maria; 106 York Drive, Princeton, NJ 08540 (US). CONNOLLY, Bernard, A.; 32 Huntingdon Close, Kingston Park, Newcastle-upon-Tyne NE2 4HH (GB).

(74) Agent: RONNING, Royal, N., Jr.; Amersham Pharmacia Biotech, Inc., 800 Centennial Avenue, Piscataway, NJ 08855-1327 (US).

(81) Designated States (national): AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— With international search report.
— Before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments.

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: IMPROVING DIDEOXYNUCLEOTIDE-TRIPHOSPHATE UTILIZATION BY THE HYPER-THERMOPHILIC DNA POLYMERASE FROM THE ARCHAEOON *PYROCOCCUS FURIOSUS*

P α -helix

Tgo 470 TIDPIEKLLDYRQRAIKILANSFYGYGYAK
Pfu 471 TQDPIEKILLDYRQKAIKLLANSFYGYGYAK

(57) Abstract: Polymerases from the Pol I family which are able to efficiently use ddNTPs have demonstrated a much improved performance when used to sequence DNA. A number of mutations have been made to the gene coding for the Pol II family DNA polymerase from the archaeon *Pyrococcus furiosus* with

the aim of improving ddNTP utilisation. "Rational" alterations to amino acids likely to be near the dNTP binding site (based on sequence homologies and structural information) did not yield the desired level of selectivity for ddNTPs. However, alteration at four positions (Q472, A486, L490 and Y497) gave rise to variants which incorporated ddNTPs better than the wild type, allowing sequencing reactions to be carried out at lowered ddNTP:dNTP ratios. Wild type Pfu-Pol required a ddNTP:dNTP ratio of 30:1; values of 5:1 (Q472H), 1:3 (L490Y), 1:5 (A486Y) and 5:1 (Y497A) were found with the four mutants; A486Y representing a 150-fold improvement over the wild type. A486, L490 and Y497 are on an α -helix that lines the dNTP binding groove, but the side chains of the three amino acids point away from this groove; Q472 is in a loop that connects this α -helix to a second long helix. None of the four amino acids can contact the dNTP directly. Therefore, the increased selectivity for ddNTPs is likely to arise from two factors: 1). small overall changes in conformation that subtly alter the nucleotide triphosphate binding site such that ddNTPs become favoured; 2) interference with a conformational change that may be critical both for the polymerisation step and discrimination between different nucleotide triphosphates.

WO 01/38546 A1

IMPROVING DIDEOXYNUCLEOTIDE-TRIPHOSPHATE UTILIZATION
BY THE HYPER-THERMOPHILIC DNA POLYMERASE FROM THE
ARCHAEON *PYROCOCCUS FURIOSUS*

5 Field of Invention

The instant disclosure pertains to DNA polymerase mutants from *Pyrococcus furiosus* which exhibit improved dideoxynucleotide utilization.

Background of Invention

10 A polymerases constitute a core component of DNA sequencing methods, a widespread and important biotechnology, based on chain-termination by dideoxynucleotide-triphosphates (ddNTPs), either the ddNTPs themselves or fluorescent derivatives. Discrimination between chain-terminating ddNTPs and dNTPs plays a key role in DNA sequencing performance. Effective ddNTP incorporation is associated with a high uniformity of signal intensity in sequencing ladders. Furthermore, efficient usage of ddNTPs requires lower concentrations; an advantage when fluorescent terminators are used, as large excesses give rise to high backgrounds. Bacteriophage T7 DNA polymerase incorporates ddNTPs much more efficiently than the enzymes from *E. coli* and *T. aquaticus* and, as a consequence, 15 gives superior sequencing ladders. The molecular basis for discrimination between dNTPs and ddNTPs resides in a single amino acid, at an equivalent location: Y526 (T7); F762 (*E. coli*); F667 (*T. aquaticus*). The T7 mutant, Y526F, shows a much reduced ability to use ddNTPs and, consequently, gives poor sequencing ladders. The F762Y and F667Y variants of *E. coli* and *T. aquaticus* use ddNTPs effectively and 20 show much improved sequencing properties. The important role of F762, in the *E. coli* polymerase, for deoxynucleotide-triphosphate recognition has been confirmed by a more complete kinetic analysis. Mutants of *T. aquaticus* DNA polymerase which have exonuclease activity removed, and contain tyrosine at position 667, have 25 excellent sequencing properties and are perhaps the most widely used enzymes for

DNA sequencing; these include Thermo-SequenaseTM and AmpliTaq FSTM DNA polymerases. All three polymerases belong to the Pol-I family (also called family A polymerases) and the critical aromatic amino acid is found in a highly conserved stretch, the B-motif (also called region III)(Figure 1). Structural data has shown that 5 the B-motif amino acids form an α -helix (the O-helix in the case of *E. coli*) with the most conserved amino acids on one side of the helix, forming part of the dNTP binding site (Figure 1). The tyrosine/phenylalanine is near the sugar of the dNTP, rationalising its critical role in dNTP/ddNTP selection.

Polymerases with thermal stability are routinely used for DNA sequencing. 10 Not only are they generally more robust than mesophilic enzymes but are essential for cycle-sequencing protocols, which involve heat-cool cycles. The extreme thermostability of polymerases purified from hyper-thermophilic archaea suggest these enzymes have potential use in DNA sequencing. However, archaeal polymerases often use ddNTPs poorly and, as a result, are generally not as useful in 15 DNA sequencing as, for example, Thermo SequenaseTM DNA polymerase. Archaeal polymerases belong to the α -family (also called the B-family, or Pol-II family), a different group to the better characterised Pol-I enzymes. However, sequence alignment shows that the Pol- α family also has a B-motif, even though it cannot be exactly aligned with that of Pol-I (8, 9) (Figure 1), and there is no exact counterpart 20 to the aromatic amino acid critical for ddNTP/dNTP selection in the Pol-I family. Nevertheless, mutations in the B-motif of α -polymerases influence dNTP binding, suggesting a role in deoxynucleoside-triphosphate recognition. Recently, crystal structures of two Pol- α members, bacteriophage RB69 gp43 and *Thermococcus gorgonarius* (*Tgo*), have been published. The B-motif amino acids form an α -helix 25 (the P-helix with both enzymes) which, as suggested by sequence alignments, is similar, but not structurally identical to the corresponding α -helix in the Pol-I family (Figure 1). Although both structures lack bound nucleic acid it was possible to model primer-template and dNTP into RB69. The B-region was located near both

primer/template and dNTP but did not appear to provide an amino acid that binds near the sugar of the dNTP. Rather, this might be supplied by tyrosine 416, an amino acid from another part of the polymerase, which packs under the sugar ring of the dNTP.

5 Sequence and structural comparisons of Pol-I and Pol- α members indicate that homologous regions are used for dNTP binding and recognition. However, the exact details of the interaction with dNTPs, and hence discrimination between dNTPs and ddNTPs, differs between the two classes. Therefore, the simple tyrosine/phenylalanine switch, so successful in converting *T. aquaticus* to a good
10 sequencing polymerase, is unlikely to be possible with archaeal polymerases.

Summary of Invention

The instant invention comprises *Pyrococcus furiosus* polymerase mutants that recognise ddNTPS a factor of up to 150 fold better and are superior for cycle-
5 sequencing protocols than the wild type enzyme. Methods for the isolation and characterisation of these mutants are discussed. The mutants exhibit superior thermal stability compared to other thermostable DNA polymerases, exhibiting stability for hours at 95°C (compared with 30-45 minutes for *T. aquaticus* polymerase), thereby permitting a greater number of cycles at elevated temperature
10 and, hence, enhanced sensitivity.

Brief Description of the Figures

The file of this patent contains at least one drawing executed in color. Copies
15 of this patent with color drawing(s) will be provided by the Patent and Trademark Office upon request and payment of the necessary fee.

FIGURE 1. A: The B-motif (region III) of polymerases from the Pol-I and Pol- α families. With Pol-I the R, K, F and YG indicated in green, and the spacing between them, are highly conserved (8-10). The F, shown in green and underlined, is
20 critical in discrimination between dNTPs and ddNTPs (4). In the case of Pol- α the Q, K, N and YG shown in green, and their spacing, are conserved. It is difficult to deduce the optimal line up, between the two families, due to variation in both the conserved amino acids and their spacing (8-10). The alignment shown (others are possible by including gaps) has the critical F in Pol-I replaced by an N (underlined)
25 in Pol- α .

FIGURE 1. B: Structure of the B-motif of a Pol-I enzyme; the *E. Coli* Klenow fragment (11). The conserved amino acids (side chains shown in green) lie on one side of an α -helix and interact with dNTP (shown in red). F762 is near the sugar ring explaining its role in dNTP/ddNTP selection.

FIGURE 1. C: Structure of the B-motif of the Pol- α enzyme from *Thermococcus gorgonarius* (28). Although the conserved amino acids (green) are on one side of an α -helix, as for Pol-I, the side chains, and their relative disposition, differ between the two families and are unlikely to interact with dNTPs identically.

5 The α -helix of Tgo-Pol is also distorted by the presence of a short stretch of β_{10} -helix between N491 and Y494, resulting in a further difference between the two families. All structures were generated using RasMol (35).

FIGURE 2. A: The amino acids in, and surrounding, the B-region of *Thermococcus gorgonarius* (Tgo) and *Pyrococcus furiosus* (Pfu) polymerases. The

10 amino acids shown in green are highly conserved and correspond to the amino acids also illustrated in green in Figure 1. Amino acids shown in magenta, when mutated in Pfu-Pol, give better sequencing performance. Note: Pfu-Pol contains a single amino acid insertion, upstream of this region, when compared to Tgo-Pol. This accounts for the numbering of corresponding amino acids differing by one.

15 FIGURE 2. B: The structure formed by the Pfu-Pol amino acids shown in part A, which comprises an α -helix (the P-helix) flanked by loop regions (shown in blue). The conserved, green, amino acids (Q484, K488, N492 and Y495) lie on one side of the helix and form part of the dNTP binding site. Three of the "improving", magenta, amino acids (A486, L490 and Y497) lie in the helical region but on the opposite side 20 to the conserved amino acids. The fourth (Q472) is in a stretch of loop that precedes the α -helix. The side chains of these amino acids are shown in space-filling mode.

25 FIGURE 2. C: End-on view of the P-helix region of Pfu-Pol clearly showing that the side chains (shown in space-filling mode) of the conserved (green) and "improving" (magenta) amino acids protrude on opposite sides of the helix. All structures were generated using RasMol (35).

FIGURE 3 is a sequencing gel in which the dideoxynucleotide-triphosphate utilization of wild type DNA polymerase from *Pyrococcus furiosus* is compared

against DNA polymerase from *Pyrococcus furiosus* mutants Q472H/N492H, A486Y, L490W, and Y497A.

FIGURE 4 is the nucleotide sequence corresponding to *Pfu* pol *exo*⁻.

5 **Detailed Description of Invention**

The following illustrates certain preferred embodiment of the invention, but is not intended to be illustrative of all embodiments,

Methods and Materials

10 *Construction of a Pfu-Pol expression vector*

A gene coding for *Pfu-Pol* (Figure 4) was digested from the plasmid pTM121 (prepared in house by Amersham Pharmacia Biotech Inc.) as a *NdeI-SmaI* fragment and ligated into *NdeI-EcoRV* cut pET-17b (Novagen) to give pET-17b(*Pfu-Pol*). This manipulation destroys the unique *EcoRV* restriction site present in pET-17b.

15 The *Pfu-Pol* gene used contains the mutation D215A, which abolishes the 3' to 5' exonuclease activity of the enzyme. All experiments were performed with the *exo*⁻ form of the enzyme.

Expression and Purification of Pfu-Pol

20 *E. coli* BL21(DE3) (Novagen) containing pET-17b(*Pfu-Pol*) was grown at 37°C until an *A*₆₀₀ of 0.5 was reached. Protein expression was induced by adding IPTG, to a final concentration of 1 mM, and continuing growth for another 4 hours. Cells were harvested by centrifugation at 4°C (5,000 rpm for 20 minutes) and resuspended in 10mM Tris pH 8.0, 100mM NaCl, 1mM PMSF, 1mM benzamidine. After sonication on ice (10 x 15s pulses) samples were centrifuged at 10,000 rpm for 20 minutes. The 25 supernatant was incubated with approximately 20 units of DNase I (Boehringer-Mannheim) for 30 minutes to hydrolyse DNA. Next the supernatant was heated at 75°C for 20 minutes to denature most of the *E. coli* proteins and inactivate the DNase I. Precipitated proteins were removed by centrifugation (10,000 rpm for 20 minutes) and the supernatant was loaded onto a 20 ml DEAE-Sephacel column, equilibrated

and eluted with Tris pH 8.0, 100mM NaCl. The flow through was collected and immediately applied to a 20 ml Heparin-Sepharose column, equilibrated with Tris pH 8.0, 100mM NaCl. The column was developed with a 100 mM - 700 mM NaCl gradient in Tris pH 8.0. Fractions were analysed by SDS-PAGE using Coomassie-
5 Blue staining and those containing a protein running at approximately 90 kD were pooled and concentrated using Centriprep 50 spin concentrators (Amicon). Protein samples were estimated to be >95% pure as judged by SDS-PAGE. All purified protein samples were stored at -20°C as 50% glycerol stocks.

Random mutagenesis of Pfu-Pol and selection of mutants by colony screening

10 Random mutagenesis was carried out on a section of the *Pfu-Pol* gene (in pET-17b) flanked by unique *EcoRV* and *SacI* restriction sites; a region comprising bases 1293-1596 and amino acid 431-532 (i.e. the P-helix and surrounding amino acids). A PCR-based method in conjunction with the mutagenic base analogues 8-oxo-dGTP and dPTP was used. The mutagenic primers used were GAACTATGATATCGCTCC
15 (*EcoRV* primer) and CTTTCAGCTCCTCCATACT (*SacI* primer). Initially 30 rounds of PCR were carried out under the following conditions: 10 mM Tris-HCl, pH 8.8, 50 mM KCl, 1.5 mM MgCl₂, 0.08% Nonidet P-40, 500 μM each dNTP (four normal and two mutagenic), 0.5 units of *T. aquaticus* DNA polymerase; cycle 94 °C - 1 min., 55 °C - 3 min., 72 °C - 2 min. The amplified products were used in a second
20 PCR reaction (conditions identical to the first but with only the four normal dNTPs at 250 μM) to generate a library of mutated DNA fragments. The library was digested with *EcoRV* and *SacI* and ligated into pET-17b(*Pfu-Pol*) from which the *EcoRV-SacI* fragment of the wild-type *Pfu-Pol* gene had been excised. The resulting plasmids were used to transform *E. coli* BL21(DE3) to ampicillin resistance.
25 Transformants containing *Pfu-Pol* variants better able to incorporate ddNTPs were selected by modifying a "colony screening rapid filter assay" usually used to detect DNA polymerase activity. Bacterial colonies were gridded onto duplicate LB/ampicillin agar plates and allowed to grow overnight at 37 °C. The colonies from

one plate were replica-plated onto a nitrocellulose filter containing activated calf-thymus DNA, which was then overlaid onto an LB/ampicillin plate containing 1 mM IPTG. After a further 4 hours, at 37 °C, during which Pfu-Pol expression was induced, the nitrocellulose filters were removed and cells lysed using 5 toluene/chloroform. The nitrocellulose filters were soaked in 20 mM Tris-HCl, pH 8.5, 10 mM KCl, 20 mM (NH₄)₂SO₄, 2 mM MgSO₄, 0.1 mg/ml bovine serum albumin and 0.1 % Nonidet P-40 and baked at 70 °C for 30 min. to destroy *E. coli* DNA polymerases. The filters were re-soaked with this buffer but also containing 12 µM each of dTTP, dCTP and dGTP plus 1 µl of [α -³²P]-ddATP (3000 Ci/mmol, 10 Nycomed-Amersham) and then incubated at 70 °C for 30 min. to allow ³²P incorporation into polymeric material. The filters were washed with trichloroacetic acid and pyrophosphate, dried and any radioactivity retained on the filters (i.e. incorporated into polymeric material) determined using autoradiography (30). Under these conditions wild type Pfu-Pol results in minimal retention of ³²P on the filters. 15 Clones that were associated with ³²P retention were rescued from the duplicate LB/ampicillin plate and used for both the preparation of mutant Pfu-Pol (exactly as for the wild type enzyme) and plasmid preparation for DNA sequencing.

Site-directed mutagenesis of Pfu-Pol

Site-directed mutagenesis of the *Pfu-Pol* gene was carried out using the PCR-based 20 “overlap extension” method (31). Most of the directed mutants were made to the P-helix i.e. within the *EcoRV/SacI* fragment described above. Therefore the *EcoRV* and *SacI* oligonucleotides were used as the common outer primers, along with appropriate primers containing the required mutation, produce two overlapping DNA fragments. These two fragments, containing the desired mutation, were used as a 25 template in a subsequent PCR reaction, together with the *EcoRV* and *SacI* primers to generate the “full-length” mutated *EcoRV/SacI* fragment. The protocol for PCR and the subsequent cloning of the mutated fragment were as described above. The amino acid Y410 is not in the P-helix and so does not lie between the *EcoRV* and *SacI*

restriction sites. The mutations Y410F/A were produced analogously using an *Eco*RI flanking primer (GGGAAAGAATTCCCTTCC) and the *Sac*I primer. The Y410 codon is located between the unique *Eco*RI and *Sac*I restriction sites.

DNA sequencing using Pfu-Pol mutants

5 DNA sequencing reactions were performed using cycle-sequencing. A kit purchased from Stratagene (Exo-“Cyclist” DNA sequencing kit) together with M13 mp18 single stranded template DNA and a universal primer was used. Approximately 1-2 units of mutant polymerase was added per sequencing reaction, corresponding to the amount of wild type Pfu-Pol normally used. Reactions were initially performed at a
10 30:1 ddNTP/dNTP ratio, the optimal nucleotide ratio for the wild type enzyme. This ratio was progressively lowered for mutant enzymes that showed an increased preference for dideoxynucleotides.

Results

Although the B-region in polymerases belonging to the Pol-I and Pol- α families are
15 similar, rather than identical, it is clear that both play a role in dNTP binding. The amino acid sequence in the vicinity of the B-region for *Pyrococcus furiosus* polymerase (Pfu-Pol), the enzyme investigated in this study, is shown in Figure 2. For comparison the homologous sequence of the structurally characterised *Thermococcus gorgonarius* enzyme (Tgo-Pol) is also shown. With Tgo-Pol The B-
20 motif consists of an α -helix (The P-helix) flanked by loop regions (Figure 2). Tgo-Pol and Pfu-Pol have 79 % sequence identity and variant amino acids invariably involve highly conservative changes. We have used Swiss-Model/SwissPdb viewer to deduce a structure for Pfu-Pol based on Tgo-Pol. As expected the two structures are almost identical and the derived P-helix from Pfu-Pol is shown in Figure 2. A
25 variety of mutations has been made to these amino acids in Pfu-Pol (Table 1), using either random mutagenesis or back-to-back PCR site-directed mutagenesis (31). The mutants were expressed in *E. coli*, using pET-17b, and purified using a heat step, followed by two chromatography columns. All proteins appeared pure by heavily-loaded SDS-PAGE, stained with Coomassie-Blue.

The DNA-sequencing performance of the mutants was assessed using a standard cycle-sequencing protocol with the four $[\alpha\text{-}^{32}\text{P}]\text{-ddNTPs}$ as chain terminators. In our hands wild-type Pfu-Pol gave good sequencing ladders (clearly defined bands, lack of non-specific termination, sequencing information spread over the entire size range of fragments) at a ddNTP:dNTP ratio of 30:1 (Figure 3). However, when a 5:1 ratio was used much of the radioactivity was found towards the top of the gel (due to lower than optimal levels of termination giving rise to long products) and excessive non-specific termination was also seen, resulting in shadow bands across all four lanes. Both effects make it impossible to deduce the DNA sequences from the sequencing ladders. As improved DNA-sequencing performance correlates with increased selectivity for ddNTPs, we looked for mutants that give a readable sequencing ladder at ddNTP:dNTP ratios of < 30:1. Therefore, mutants were initially tested at the wild-type 30:1 ratio. Enzymes which gave readable sequencing gels or were associated with increased radioactivity towards the bottom of the gel (indicative of better ddNTP incorporation resulting in shorter products) were then evaluated at a 5:1 ratio of ddNTP:dNTP and, if warranted, at progressively decreasing ratios of the two triphosphates.

Initially, the four most highly conserved amino acids (Q484, K488, N492, Y495), on the side of the P-helix that faces the ddNTP binding site (Figures 1 and 2), were investigated. Most of the mutants showed inferior sequencing ladders at the 30:1 ddNTP:dNTP ratio used for initial screening, when compared to the wild type enzyme (Table 1). The few mutants that appeared equivalent to the wild type under these conditions all gave unreadable sequencing gels, characterised by radioactivity towards the top of the gel and non-specific termination, at a 5:1 ratio of trinucleotides. Both sequence and structural comparisons (Figure 1) show that the Pol- α family does not have an exact counterpart to the critical phe/tyr found with Pol-I enzymes. The best guess as to which amino acid, if any, would play this role is N492; a residue that is highly conserved (8, 9) and structurally at least in a similar location to the phe/tyr (Figure 1). However, N492Y (i.e. an attempt to introduce a tyr

near a location shown to give good sequencing performance with the Pol-I family) gave an inactive enzyme. Other mutations to this amino acid, e.g. N492H/K/G, while active resulted in difficult to read ladders at the initial 30:1 ddNTP:dNTP ratio. Changes to Q484, K488 and Y495 gave sequencing phenotypes at best equivalent to, 5 but more commonly worse than, the wild type enzyme (Table 1).

As shown in Figure 1, not only do the conserved amino acids on the same side of the O- or P-helix vary in the nature of their side chains but also in the spacing between them. The B-motif in the Pol- α family is missing an amino acid when compared to Pol-I. Alternative alignments to that shown in Figure 1, in which a gap 10 in Pol- α is placed opposite the key phe/tyr in Pol-I, have been proposed (8, 9). Therefore, insertion mutations consisting of: addition of single tyrosine immediately before N492 (A(Y)N); replacement of several Pfu-Pol P-helix amino acids with the corresponding regions of the *T. aquaticus* O-helix (TIN(Y)GVL) (Table 1), have been prepared. In these Pfu-Pol variants the "missing" amino acid is replaced with a 15 tyrosine either in the Pfu-Pol or the *T. aquaticus* polymerase context, allowing a more exact sequence alignment of the mutated B-motif with that of the Pol-I family. These insertion mutations also represent another approach to placing a tyrosine near the important phe/tyr in the pol-I family. Unfortunately, all insertion mutations were inactive (Table 1).

20 Following alterations to the conserved amino acids and variation in their spacing a number of mutations have also been made in most of the other P-helix amino acids. As can be seen in Table 1 the majority of the changes lead to a sequence performance roughly equivalent to the wild type polymerase i.e. readable ladders at ddNTP:dNTP ratios of 30:1, no useful sequencing data, with most radioactivity at the 25 top of the gel and non-specific termination at the 5:1 ratio. Therefore, most changes offered no improvement in DNA sequencing performance over that of the wild type. Three mutations, A486Y, L490W and Y497A, were found which gave readable sequencing ladders at the 5:1 ratio. Progressively decreasing the amount of chain terminator showed that A486Y could be used at a ddNTP:dNTP ratio of 1:5, L490W

at 1:3 and Y497A at 5:1; representing an improvement over wild type (in terms of ddNTP usage) of 150, 90 and 6 fold respectively (Table 1, Figure 3). Other mutations to these amino acids e.g. L490Y and Y497W also resulted in preferential usage of ddNTPs when compared to the wild type, although the ddNTP:dNTP ratios useable with these mutants were not as good as with L490W and Y497A respectively. We also found certain mutations to these amino acids e.g. A486W and Y497F (Table 1) that resulted in poorer ddNTP utilisation as compared to the wild type enzyme.

As well as changing the amino acids in the P-helix in a directed manner, random mutagenesis has been used between amino acids 431 and 532. This stretch of 101 amino acids encompasses the P-helix as well as amino acids flanking this structural element (Figure 2). Preparing large number of random mutants is quick and straightforward; screening for improved phenotype is often more tedious and difficult. We have adapted a "colony screening rapid filter assay", originally used to identify recombinants containing wild type *T. aquaticus* DNA polymerase, to directly assess improved incorporation of ddNTPs. Essentially *E. coli* colonies, on nitrocellulose filters containing activated calf-thymus DNA, are incubated with [α -³²P]-ddATP and dTTP/dCTP/dGTP. Colonies expressing a Pfu-Pol mutant that effectively uses ddNTPs transfer radioactivity to polymeric DNA immobilised on the nitrocellulose filter, allowing subsequent detection by autoradiography. The method includes a heat step, prior to the assay, to destroy host polymerases and so automatically scores for thermostability in the mutant Pfu-Pol. Only one mutant, that appeared to incorporate ddNTPs particularly well, was found and revealed by sequencing to be a double mutant, Q472H-N492H. This double mutant gave good sequencing data at a ddNTP:dNTP ratio of 5:1 (Table 1) i.e. a 5-fold improvement over wild type. However, the amount of ddNTP could not be further decreased and so the double mutant does not appear to be as good as either A486Y or L490W. The amino acid at 492, normally asparagine, is one of the conserved P-helix amino acids, described above (Figures 1 and 2). We therefore thought that the improvement in phenotype was due to the N492H change, a combination not tested above.

Remarkably single directed mutants showed that the key change was to Q472H, which gave good sequencing ladders, equivalent to those produced by the double mutant, at 5:1 ddNTP:dNTP ratios. The single mutant N492H showed non-specific termination at the 5:1 ratio and in fact gave slightly inferior performance to the wild type when tested at 30:1 ddNTP:dNTP. The amino acid Q472 is located in a loop that connects the P-helix to another long α -helix (Figure 2) and its location is such that it is unlikely to interact directly with dNTPs. Not only does the mutation Q472H improve ddNTP incorporation but it fully rescues a mutant, N492H, with sequencing performance worse than the wild type.

Changes have also been made to Y410, an amino acid that is not in the P-helix, but has been shown by crystallography to have an important role in binding dNTPs, forming the bottom of the nucleotide binding cleft. In our hands both Y410F and Y410A offered no improvement over the wild type (Table 1).

Attempts to further reduce the amount of ddNTP needed for sequencing, by combining the improved mutants A486Y, L490W and Q472H-N492H were unsuccessful. One double mutant A486Y-L490W behaved in an identical manner to the single mutant A486Y i.e. gave readable sequencing gels at 1:5 ddNTP:dNTP (Table 1). The other two, A486Y-Q472H-N492H and L490W-Q472H-N492H, actually gave worse sequencing performance, characterised by unreadable gels at ddNTP:dNTP ratios of 1:5 and 1:3, than was seen with A486 and L490W alone.

Thus, it has been seen that with the Pol-I enzymes a change of a single phe to a tyr converts ddNTPs from very poor to very good substrates (4). It was suggested that the 3'-OH of the natural dNTP substrate was one of the ligands for the essential Mg^{2+} , required for polymerisation. With ddNTPs, which lack a 3'-OH, this interaction cannot take place, accounting for their poor substrate properties. Most Pol-I enzymes contain a phe, near the sugar of the dNTP, which is sometimes (either in a few natural Pol-I enzymes e.g. from T7 or by site-directed mutagenesis) replaced by tyr.

Further, changes to most of the amino acids likely to interact directly with the dNTP do not improve discrimination for ddNTPs over dNTPs. This includes the conserved amino acids on one side of the P-helix and Y410 (Figures 1 and 2, Table 1). None of the mutations improve ddNTP incorporation and sequencing performance (or else give low specific activities rendering the question of improved properties irrelevant) and an effect as profound as that caused by the single phe/try switch with the Pol-I family was never observed. Additionally, inserting a single amino acid into the P-helix of Pfu-Pol (to produce a better alignment with the O-helix of Pol-I) or helix swaps between the two categories gave inactive enzyme. This probably arises because of a large overall disruption to Pol- α structure when its P-helix is replaced with the equivalent of an O-helix. Finally, changes to most of the other, less conserved, amino acids in the P-helix also lead to no improvement in selectivity for ddNTP and DNA sequencing performance (Table 1).

Changes at three positions within the P-helix, amino acids A486, L490 and Y497 gives rise to a higher selectivity for ddNTPs. As shown in Figure 2, the side chains of these amino acids protrude from the P-helix on the side facing away from the DNA-binding cleft. The three amino acids cannot, therefore, interact directly with the dNTP, as postulated for the four highly conserved amino acids, Q484, K488, N492 nad Y495. Mutagenesis of VentTM polymerase also concluded that changes to the equivalent of the Pfu-Pol A486 (A488 in VentTM) lead to an increase in the efficiency of ddNTP usage (33). This study changed A488 to several alternatives and it was observed that the bigger the side chain, the better the utilisation of ddNTP. The largest side chain used in this study was phe and this lead to a boost in the incorporation of ddNTPs, by a factor of about 15-fold, evaluated using sequencing gels. The change we made, A486Y, allows readable sequencing gels at a ddNTP:dNTP ratio of 1:5, a 150-fold improvement over the wild type. The determination of "sequencing performance" by visual inspection of gels is somewhat subjective; even under optimised conditions with the wild-type enzyme, non-specific termination is visible in a few places. Therefore, deciding the ddNTP:dNTP ratio

which gives clearly interpretable patterns i.e. exactly when non-specific termination begins to interfere with readability is difficult. However, we routinely use the Pfu-Pol A486Y mutant for sequencing purposes at the 1:5 ddNTP:dNTP ratio and find it gives reliable sequencing information. What is clear, though, is that this amino acid 5 represents a residue important for ddNTP usage, even though any improvement is far less than the phe/tyr switch in the pol-I family. L490 is one turn along the helix form A486 and the side chains of these two amino acids point in almost identical directions (Figure 2). This amino acid was not studied with VentTM but we have found that the mutation L490W improves ddNTP usage, by a factor only slightly less. 10 than the A486Y change. L490Y, while slightly improved over wild type, is far less effective than L490W. This may, like the changes to the alanine described above, result from a correlation between side chain size and ddNTP utilisation. However, as there are not many larger side chains than the wild type leu, it is difficult to test this 15 experimentally with a series of mutations as carried out with VentTM polymerase and A488 (32). Two other P-helix amino acids, L479 and Y497, have side chains that point in almost the same direction as those of A486 and L490. Changes to L479 give a phenotype similar to wild type (Table 1). However, some changes to Y497 (Y497W and Y497A) lead to a slightly improved preference for ddNTPs (Table 1), whereas others Y497F decrease ddNTP utilisation. A similar effect was seen with 20 VentTM with the corresponding amino acid Y499 (33).

Not only are A488 and L490, the amino acids in the P-helix which when changed lead to the most pronounced preference for ddNTPs, unable to interact with dNTPs directly, but they are not highly conserved in the Pol- α family. Although the equivalent position to A486 is most usually ala, other amino acids e.g. asn, ser, ile, 25 leu and phe are found (8, 9). Similarly L490 is commonly replaced by either another hydrophobic amino acid or thr/ser. This non-conservation emphasises that A486 and L490 are unlikely to have a direct critical function e.g. binding dNTP, primer/template or in catalysis.

The indirect action of the mutants A488Y and L490W was confirmed by the detection of a third mutant, associated with better ddNTP usage. Random mutagenesis to a region including both the P-helix and its flanking sequences revealed the double mutant Q472H-N492H. The improved properties do not result 5 from the change to N492 (the conserved N in the P-helix, Figures 1 and 2) but, as shown by subsequent single mutants (Table 1), to alteration to Q472. The residue Q472 is in a loop that lies between the P-helix and a second long α -helix. These two anti-parallel helices comprise the bulk of the fingers domain of the RB69 (27) and *Tgo* (28) polymerases and presumably also of the Pfu-Pol. Q472 is highly variable 10 between Pol- α enzymes and even differs within the more closely related archaeal sub-set of this class (8, 9) e.g. this residue is ile with *Tgo*-Pol and VentTM (Figure 2). Q472H, or the double mutant Q472H-N492H, only improve ddNTP incorporation, as assessed by sequencing gels, by a factor of 6-fold.

It is apparent that many modifications and variations of the invention as 15 hereinabove set forth may be made without departing from the spirit and scope thereof. The specific embodiments described are given by way of example only, and the invention is limited only by the terms of the appended claims.

31. Table 1. DNA sequencing properties of Pfu-Pol mutants.

Category of enzyme	Mutant	ddNTP:dNTP ratio optimized for DNA sequencing ladders
Wild-type	Wild-type	30:1
Alterations to highly conserved amino acids in P-helix (shown in green in Figures 1 and 2)	Q484A K488A N492H/K/G N492Y Y495/I/D/C/S	Worse than wild type Worse than wild-type Worse than wild-type Inactive Enzyme Similar to wild-type
Insertions into P-helix*	A(Y)N and TIN(Y)GVL	Inactive enzyme
Alterations to other amino acids in P-helix (see Figure 2)	L479Y/W/P A486Y A486W L490W L490Y S493Y F494Y/C/S/T/V G496P/S/A Y497F Y497W Y497A	Similar to wild-type 1:5 Worse than wild-type 1:3 5:1 Similar to wild-type Similar to wild-type Similar to wild-type Worse than wild-type 10:1 5:1
Alterations to Y410	Y410A/F	Worse than wild-type
Alterations to loop preceding the P-helix (see Figure 2)	N492H	5:1
Multiple mutations	Q472H-N492H A486Y-L490W Q472H-A486Y-N492H Q472H-L490W-N492H	5:1 1:5 Worse than A486Y Worse than L490W

5 The wild-type enzyme gives readable DNA sequencing gels at a 30:1 ddNTP:dNTP ratio, but unreadable gels at 5:1. Mutants described as similar to wild-type have this property; mutants noted as worse give

unreadable gels at the 30:1 ratio. For mutants with improved discrimination the number given is the lowest ratio of ddNTP:dNTP at which readable sequencing gels were obtained. Y410F/A indicates Y 410 was changed to both F and A etc. *A(Y)N has a Y inserted between A491 and N492. In TIN(Y)GVL this sequence replaces the amino acids between 489 and 494 (LLANSF).

What is claimed is:

- 5 1. A purified recombinant thermostable DNA polymerase comprising the amino acid sequence which corresponds to the nucleotide sequence set forth in Figure 4 further modified to contain one or more amino acid changes selected from the group consisting of Q472H/N492H, A486Y, L490W, L490Y, Y497A, Y497W, N492H, Q472H/N492H, and A486Y/L490Y.
- 10 2. An isolated nucleic acid that encodes a thermostable DNA polymerase having the peptide sequence of Claim 1.
3. A DNA polymerase encoded for by the nucleic acid of Claim 2.
4. A recombinant DNA vector that comprises the nucleic acid of Claim 2.
5. The recombinant vector of Claim 3 containing the plasmid pTM121.
- 15 6. A recombinant host cell transformed with the vector of Claim 4.
7. The recombinant host cell of Claim 5, wherein the cell is *E. coli*.
8. In a method of sequencing DNA which comprises the step of generating chain terminated fragments from a DNA template to be sequenced with a DNA polymerase, the improvement comprising utilizing, as said DNA polymerase, the DNA polymerase of Claim 2 in the presence of at least one chain terminating agent and one or more nucleotide triphosphates, and determining the sequence of said DNA from the sizes of said fragments.
- 20 9. A kit for sequencing DNA comprising the DNA polymerase of of Claims 2 .

1/5

Fig. 1A

Pol I $\cdots (X)_3 \cdots (X)_3 \cdots (X)_3 \cdots$
Pol α $\cdots (X)_3 \cdots (X)_3 \cdots (X)_2 \cdots$

Fig. 1B

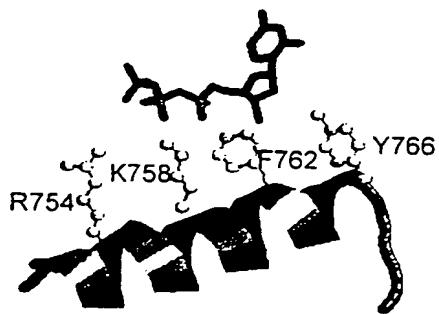
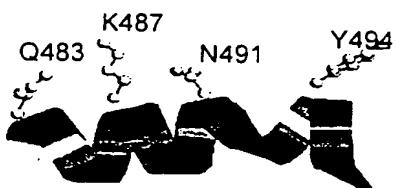


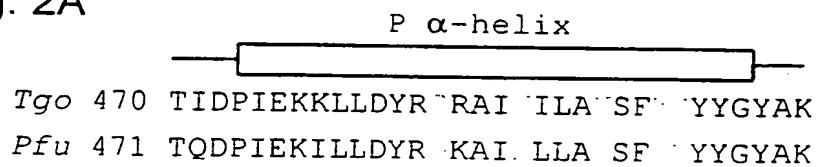
Fig. 1C



2/5

Fig. 2A

P α -helix



Tgo 470 TIDPIEKLLDYR RAI ILA SF YYGYAK
Pfu 471 TQDPIEKILLDYR KAI LLA SF YYGYAK



Fig. 2B

Fig. 2C

3/5

WT WT ^{Fig. 3} Q472H/ A486Y L490W Y497A
N492H

30:1 5:1 5:1 1:5 1:3 5:1

A G C T A G C T A G C T A G C T A G C T A G C T

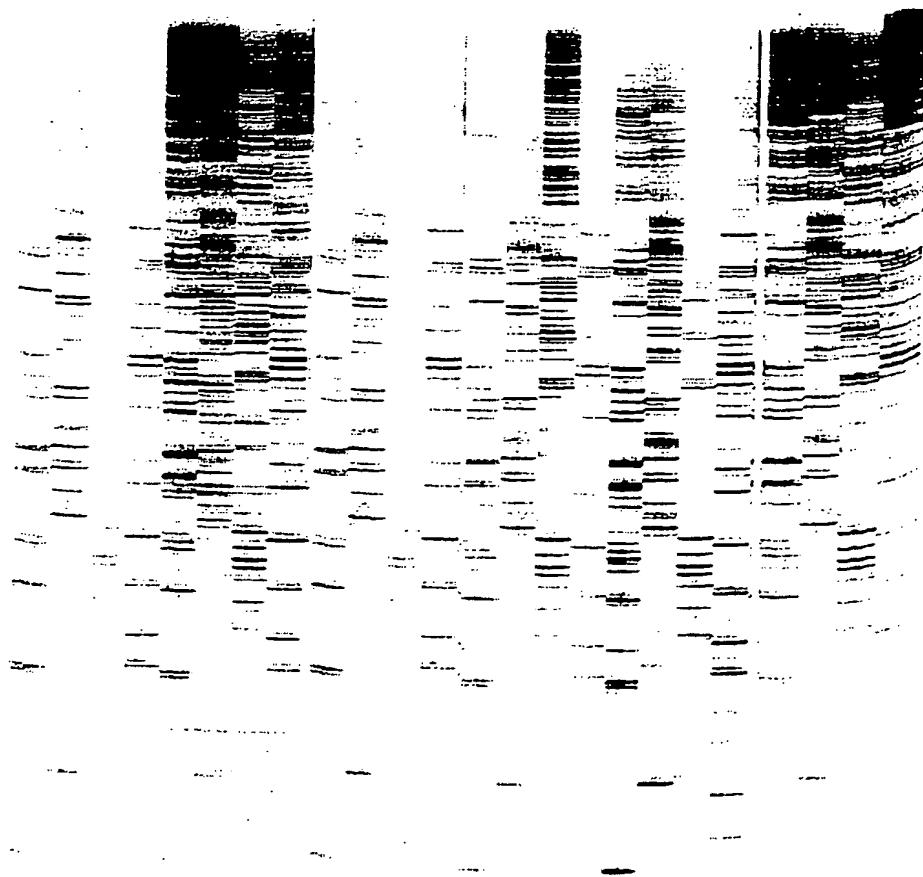


FIG. 4

Pfu pol exo⁻ DNA sequence

ATGATTTAGATGGGATTACATAACTGAAGAAGGAAACCTGTTAGGCTATTCAAAAGAGAACG TACTAAATCTACACCTAATGTATTGACTTCTTCAGGGATGATC Base pairs 1-70
AAAAATTAAAGATAGAGCATGATAGAACCTTAGACCATACATTACGCTCTCAGGGATGATC CTTTAAATTCTATCTGACTATCTGAAATCTGGTATGAAATGCGAGAAGAGTCCCTACTAAG Base pairs 71-140
GATTGAAGAAGTTAAGAAAATAACGGGGAAAGGCATGGAAAGATTGTGAGAAATTGTTGATGAGAGAAG CTAACTCTTCAATTCTTATTGCCCCCTTCCGTACCTTCAACACTCTAACACTACATCTTC Base pairs 141-210
GTGAGAAAAAGTTCTCGCAAGCCTATTACCGTGTGGAAACCTTATTGGAACATCCCAAGATGTT CAACTCTTTCAAGAGCCGTCGGATAATGGCACACCTTGAATAACCTGTAGGGGTTCTACAAG Base pairs 211-280
CCACTATTAGAGAAAAGTTAGAGAACATCCAGCAGTTGTGGACATCTCGAATACGATATTCCATTG GGTGATAATCTCTTTCAATCTTGTAGTCACACCTGTAGAACCTTATGCTATAAGTAAACCG Base pairs 281-350
AAAGAGATAACCTCATCGACAAAGGCCAATACCAATGGAGGGGAAGAAGAGCTAAAGATTCTGCCTTC TTCTCTATGGAGTAGCTGTTCCGGATTATGGTACCTCCCCCTTCTCGATTCTAAGAACCGGAAG Base pairs 351-420
GATATAGAAAACCTCTATCACGAAGGAGAACAGAGTTGGAAAAGGCCAAATTATAATGATTAGTTATGCAG CTATATCTTGGAGATAGTGTCTCTTCTCAACCTTCCGGTTAATTACTAACATCAACAGTC Base pairs 421-490
ATGAAAATGAAGCAAAGGTGATTACTTGGAAAAACATAGATCTTCCATACGTTGAGGGTGTATCAAGCGA TACTTTACTTGTCTTCAACTAATGAACCTTTGTATAGAACCTGCTATAAGTATGCAACTCAACATAGTCGCT Base pairs 491-560
GAGAGAGATGATAAAAGAGATTCTCAGGATTATCAGGGAGAAGGATCCTGACATTATAGTTACTTATAAT CTCTCTACTATTTCTAAAGAGTCCTAATAGTCCCTCTCCCTAGGACTGTAAATATCAATGAATATTA Base pairs 561-630
GGAGACTCATTCGATTCCATATTTAGCAGGGCAGAAAACCTGGGATTAAATTAAACCATGGAA CCTCTGAGTAAGCGTAAGGGTATAATCGCTTTCCGTCTTTGAACCTAATTAAATTGGTAACCTT Base pairs 631-700
GAGATGGAAGCGAGGCCAAGATGCAGAGAACAGCGATATGACGGCTGTAGAACGTCAAGGGAGAACACA CTCTACCTTCGCTGGTTCTACGCTCTTACCGCTATACTGCCGACATCTCAGTCCCTCTATGT Base pairs 701-770
TTTCGACTTGTATCATGTAATAACAAGGACAATAATCTCCAACATACACACTAGAGGCTGTATATGAA AAAGCTGAACATAGTACATTATTCTCTGTATTAGGGTTGTATGTGTATCTCGACATATACTT Base pairs 771-840
GCAATTGGAAAGCCAAGGAGAACGGTATACGCCGACGAGATAGCAGGGCTGGAGAACGCTGGAAAGTGGAGAGA CGTTAAAAACCTTCGGTTCTCCATATGCCGCTCTATCGTTCCGACCTTCACCTCT Base pairs 841-910
ACCTTGAGAGAGTTGCCAATACCGATGGAAGATGCAAAGGCAACTTATGAAACTCGGGAAAGAACATT TGGAACTCTCAACGGTTATGAGCTACCTCTACGTTCCGTTGAATACTTGAGGCCCTTCTTAAGGA Base pairs 911-980
TCCAATGAAATTCAAGATTAGTTGGACAACCTTATGGGATTTCAAGGTCAAGCACAGGG AGGTTACCTTAAGTCGAAAGTTCTAATCAACCTGTGGAAATACCCCTACAAAGTCCAGTCGTGCCCC Base pairs 981-1050
AACCTGTAGAGTGGTTCTACTTAGGAAAGCCTACGAAAGAACGAAGTAGCTCCAAACAAGCCAAGTG TTGGAACATCTACCAAGAACATGAATCCCTTCGATGCTTCTTGTCTACGAGGTTGGTCCAC Base pairs 1051-1120
AAGAGGAGTATCAAAGAACGGCTAGGGAGAGCTACACAGGGATTCTAAAGAGGCCAGAACAGGG TTCTCCATAGTTCTCGAGTCCCTCTCGATGTTGCTCCACCTAACAGTCCAGTCGTGCCCC Base pairs 1121-1190
GTGGGAAACATAGTACCTAGATTAGGGCTATATCCCTCGATTATAATTACCCACAATGTTCT CACCCCTTGTATCATGGATCTAAATCTGGGATATAGGGAGCTAACATTAAATGGGTGTACAAAGA Base pairs 1191-1260
CCCGATACTCTAAATCTTGAGGGATGCAAGAACATGATATCGCTCTCAAGTAGGCCACAAGTCTGCA GGGCTATGAGATTAGAACCTCAGTTCTGATACTATAGCGAGGACTCATCCGGTGTCAAGACGT Base pairs 1261-1330
AGGACATCCCTGGTTTATACCAAGTCTTGGACATTGTTAGGGAAAGACAAAGATTAAGACAAA TCCTGTAGGGACCAAAATATGGTCAAGAGAACCTGTAAACAACTCTCTGTCTTCTAATTCTGTT Base pairs 1331-1400
AATGAAGGAAACTCAAGATCCTATAGAAAAAAACTCCCTGACTATAGACAAAAAGCGATAAAACTCTTA TTACTTCTTGTAGTTCTAGGATATCTTTTTATGAGGAACGTGATATCTGTTTCGCTATTTGAGAAAT Base pairs 1401-1470
GCAAATTCTTCTACGGTATTATGGCTATGCAAAGCAAGATGGTACTGTAAGGAGTGTGCTGAGAGCG CTTTAAGAAAGATGCCATAATACCGATACTGTTCTACCATGACATTCTCACAGCAGTCGCG Base pairs 1471-1540

FIG. 4 Continued

TTACTGCCTGGGAAGAAAAGTACATCGAGTTAGTATGAAAGGAGCTCGAAGAAAAGTTGGATTTAAAGT AATGACGGACCCCTTCTTCATGTAGCTAACATACCTTCCTCGAGCTCTTTCAACCTAAATTCA	Base pairs 1541-1610
CCTCTACATTGACACTGATGGTCTCTATGCAACTATCCCAGGAGGAGAAAGTGAGGAATAAAGAAAAG GGAGATGTAAGTGTGACTACCAAGAGATACGTTGATAGGGTCCCTCTTCACTCCTTATTTCTTTTC	Base pairs 1611-1680
GCTCTAGAATTGTAAAATACATAAAATTCAAAGCTCCCTGGACTGCTAGAGCTTGAATATGAAGGGTTT CGAGATCTTAAACATTATGTATTTAGTTAGTTCGAGGGACCTGACGATCTGAACTTATACTTCCCAGAA	Base pairs 1681-1750
ATAAGAGGGGATTCTCGTTACGAAGAAGAGGTATGCACTAATAGATGAAGAAGGAAAAGTCATTACTCG TATTCTCCCTAAGAACATGCTTCTCCATACGTCATTACTCTTCTTCAGTAATGAGC	Base pairs 1751-1820
TGGTTTAGAGATAGTTAGGAGAGATTGGAGTGAAGAAGCTGAAATTGCAAAAGAAAACCTCAAGCTAGAGTTTGGAGACA ACCAAATCTCTATCAATCCCTCTAACCTCACTTAACGTTTGTGAGTTGATCTCAAAACCTCTGT	Base pairs 1821-1890
ATACTAAACACGGAGATGTTGAGAAGAGCTGTGAGAATAGTAAAAGAAGTAATACAAAAGCTGCCAATT TATGATTTGTGCCCTACAACCTCTCGACACTCTTATCATTTCTTCATTATGTTGACCGTTAA	Base pairs 1891-1960
ATGAAATTCCACCAGAGAACGCTCGCAATATATGAGCAGATAACAAGACCAATTACATGAGTATAAGGGCAGT TACTTTAAGGTGGTCTCTCGAGCGTTATATACTCGTCTATTGTTCTGGTAATGTAACCTCATATTCCGCTA	Base pairs 1961-2030
AGGTCCCTACGTAGCTGTTGCAAAGAAAAGCTGCTAAAGGAGTTAAAATAAAGCCAGGAATGGTAATT TCCAGGAGTGCATCGACAACGTTCTTGATCGACGATTCTCAATTATTCGGTCCCTACCAATTAA	Base pairs 2031-2100
GGATACATAGTACTTAGAGGGCGATGGTCCAATTAGCAATAGGGCAATTCTAGCTGAGGAATACGATCCC CCTATGTATCATGAATCTCCGCTACCAAGGTTAACGTTATCCGTTAAGATCGACTCCTTATGCTAGGGT	Base pairs 2101-2170
AAAAGCACAAGTATGACGCAGAATATTACATTGAGAACCCAGGTTCTTCCAGCGGTACTTAGGATATTGGA TTTCGTGTTCATACTCGCTTATAATGTAACCTTGGTCCAAGAAGGTGCCATGAATCCTATAACCT	Base pairs 2171-2240
GGGATTTGGATACAGAAAAGGAAGACCTCAGATAACCAAAAGACAAGACAAGTCGGCCTAACCTCCTGGCTT CCCTAAACCTATGTCTTCTTCTGGAGTCTATGGTTCTGTTCTGAGCCGGATTGAAGGACCGAA	Base pairs 2241-2310
AACATTAACCTCTAG Base pairs TTGTAATTTTAGGATC 2311-2328	

INTERNATIONAL SEARCH REPORT

Int. Application No

PCT/US 00/31830

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/54 C12N9/12 C12N1/21 C12Q1/68 // (C12N1/21, C12R1:19)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

WPI Data, PAJ, CAB Data, STRAND, BIOSIS, EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P, X	EVANS STEVEN J ET AL: "Improving dideoxynucleotide-triphosphate utilisation by the hyper-thermophilic DNA polymerase from the archaeon Pyrococcus furiosus." NUCLEIC ACIDS RESEARCH, vol. 28, no. 5, 1 March 2000 (2000-03-01), pages 1059-1066, XP002164138 ISSN: 0305-1048 the whole document ----	1-9
X	US 5 827 716 A (MAMONE JOSEPH A) 27 October 1998 (1998-10-27) column 1, line 60 -column 28 ----	1-9
Y	---- -/-	1-9

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

* Special categories of cited documents :

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

T later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

X document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

Y document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

& document member of the same patent family

Date of the actual completion of the international search

28 March 2001

Date of mailing of the international search report

09/04/2001

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel: (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Hornig, H

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 00/31830

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	EP 0 727 496 A (HARVARD COLLEGE) 21 August 1996 (1996-08-21) page 17, line 30 -page 18, line 31; claims 1-3,5,6 page 17, line 37 - line 39 ----	1-9
Y	EP 0 655 506 A (HARVARD COLLEGE) 31 May 1995 (1995-05-31) page 18, line 45 -page 19, line 20; claims 1-4,8-18,35,47-52 ----	1-9
A	US 5 489 523 A (MATHUR ERIC J) 6 February 1996 (1996-02-06) the whole document ----	
A	WO 99 06538 A (AMERSHAM PHARM BIOTECH INC) 11 February 1999 (1999-02-11) the whole document ----	
A	GARDNER ANDREW F ET AL: "Determinants of nucleotide sugar recognition in an archaeon DNA polymerase" NUCLEIC ACIDS RESEARCH, OXFORD UNIVERSITY PRESS, SURREY, GB, vol. 27, no. 12, 15 June 1999 (1999-06-15), pages 2545-2553, XP002163101 ISSN: 0305-1048 the whole document -----	

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 00/31830

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
US 5827716	A	27-10-1998	NONE	
EP 0727496	A	21-08-1996	US 5614365 A DE 29513622 U DE 29513639 U JP 2691149 B JP 8205874 A AT 143057 T AU 4193396 A CA 2201885 A CN 1170439 A DE 69400567 D DE 69400567 T DE 655506 T DK 655506 T EP 0655506 A ES 2072238 T FI 971611 A GR 95300040 T HU 77037 A WO 9612042 A ZA 9508761 A	25-03-1997 19-10-1995 19-10-1995 17-12-1997 13-08-1996 15-10-1996 06-05-1996 25-04-1996 14-01-1998 24-10-1996 06-02-1997 28-09-1995 10-03-1997 31-05-1995 16-07-1995 13-06-1997 31-07-1995 02-03-1998 25-04-1996 13-05-1996
EP 0655506	A	31-05-1995	US 5614365 A AT 143057 T AU 4193396 A CA 2201885 A CN 1170439 A DE 29513622 U DE 29513639 U DE 69400567 D DE 69400567 T DE 655506 T DK 655506 T EP 0727496 A ES 2072238 T FI 971611 A GR 95300040 T HU 77037 A JP 2691149 B JP 8205874 A WO 9612042 A ZA 9508761 A	25-03-1997 15-10-1996 06-05-1996 25-04-1996 14-01-1998 19-10-1995 19-10-1995 24-10-1996 06-02-1997 28-09-1995 10-03-1997 21-08-1996 16-07-1995 13-06-1997 31-07-1995 02-03-1998 17-12-1997 13-08-1996 25-04-1996 13-05-1996
US 5489523	A	06-02-1996	US 5948663 A US 5545552 A WO 9209689 A US 5866395 A	07-09-1999 13-08-1996 11-06-1992 02-02-1999
WO 9906538	A	11-02-1999	US 5882904 A AU 8767198 A EP 1002057 A	16-03-1999 22-02-1999 24-05-2000